RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY SYSTEMS BRANCH
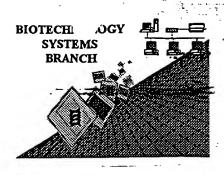
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: *09/486,167*

Source: *1643*

Date Processed by STIC: *10-03-00*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,

2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR FURTHER INFORMATION, PLEASE TELEPHONE MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

N/A   1643

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/486,167*

DATE: 10/03/2000
TIME: 22:46:46

INPUT SET: S35961.raw RECEIVED

OCT 03 2000

TECH CENTER 1600/2900

This Raw Listing contains the General
Information Section and up to the first 5 pages.

Does Not Comply
Corrected Diskette Needed

```
 1                          SEQUENCE LISTING
 2
 3    (1)    General Information:
 4                                      Applicant name here
-->  5        (i) APPLICANT:
-->  6            (A) NAME: UNIVERSITE CATHOLIQUE DE LOUVAIN
 7                       Halles Universitaires
-->  8            (B) STREET: Place de l' Universite, 1
-->  9            (C) CITY: LOUVAIN-LA-NEUVE
--> 10            (E) COUNTRY: BELGIUM
--> 11            (F) POSTAL CODE (ZIP): B-1348
 12
--> 13            (A) NAME: UNIVERSITE DE MONS-HAINAUT
--> 14            (B) STREET: Place du Parc 20
--> 15   (IV)     (C) CITY: MONS
--> 16            (E) COUNTRY: BELGIUM
--> 17            (F) POSTAL CODE (ZIP): B-7000
 18
 19        (ii) TITLE OF INVENTION: PEROXISOME-ASSOCIATED PEPTIDE, NUCLEOTIDE
 20             SEQUENCE ENCODING SAID PEPTIDE AND THEIR USES IN THE
 21             DIAGNOSTIC AND/OR THE TREATMENT OF LUNG INJURIES AND
 22             DISEASES, AND OF OXIDATIVE STRESS-RELATED DISORDERS
 23
 24        (iii) NUMBER OF SEQUENCES: 19
 25
 26   (V)(iv) COMPUTER READABLE FORM:
 27            (A) MEDIUM TYPE: Floppy disk
 28            (B) COMPUTER: IBM PC compatible
 29            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 30            (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 31
 32
 33    (2) INFORMATION FOR SEQ ID NO: 1:
 34
 35        (i) SEQUENCE CHARACTERISTICS:
 36            (A) LENGTH: 805 base pairs
 37            (B) TYPE: nucleic acid
 38            (C) STRANDEDNESS: single
 39            (D) TOPOLOGY: linear
 40
 41        (ii) MOLECULE TYPE: cDNA
 42
 43        (iii) HYPOTHETICAL: NO
 44
 45        (iv) ANTI-SENSE: NO
 46
```

(i) GENERAL INFORMATION:
  (i) APPLICANT:
  (ii) TITLE OF INVENTION:
  (iii) NUMBER OF SEQUENCES:
  (iv) CORRESPONDENCE ADDRESS:
    (A) ADDRESSEE:
    (B) STREET:
    (C) CITY:
    (D) STATE:
    (E) COUNTRY:
    (F) ZIP:

(v) COMPUTER READABLE FORM:
  (A) MEDIUM TYPE:
  (B) COMPUTER:
  (C) OPERATING SYSTEM:
  (D) SOFTWARE:
(vi) CURRENT APPLICATION DATA:
  (A) APPLICATION NUMBER:
  (B) FILING DATE:
  (C) CLASSIFICATION:

(VI) Current Application Data missing.

EPO format invalid for U.S. applications.

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/486,167*

INPUT SET: S35961.raw

```
47       (vi) ORIGINAL SOURCE:
48            (A) ORGANISM: Homo sapiens
49
50      (ix) FEATURE:
51            (A) NAME/KEY: CDS
52            (B) LOCATION:193..681
53
54
55      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
56
57   GCCAGGAGGC GGAGTGGAAG TGGCCGTGGG GCGGGTATGG GACTAGCTGG CGTGTGCGC     60
58
59   CTGAGACGCT CAGCGGGCTA TATACTCGTC GGTGGGGCCG GCGGTCAGTC TGCGGCAGCG    120
60
61   GCAGCAAGAC GGTGCAGTGA AGGAGAGTGG GCGTCTGGCG GGGTCCGCAG TTTCAGCAGA    180
62
63   GCCGCTGCAG CC ATG GCC CCA ATC AAG GTG GGA GAT GCC ATC CCA GCA        228
64              Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ala
65               1           5                   10
66
67   GTG GAG GTG TTT GAA GGG GAG CCA GGG AAC AAG GTG AAC CTG GCA GAG     276
68   Val Glu Val Phe Glu Gly Glu Pro Gly Asn Lys Val Asn Leu Ala Glu
69           15                  20                  25
70
71   CTG TTC AAG GGC AAG AAG GGT GTG CTG TTT GGA GTT CCT GGG GCC TTC     324
72   Leu Phe Lys Gly Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe
73       30                  35                  40
74
75   ACC CCT GGA TGT TCC AAG ACA CAC CTG CCA GGG TTT GTG GAG CAG GCT     372
76   Thr Pro Gly Cys Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala
77    45                  50                  55                  60
78
79   GAG GCT CTG AAG GCC AAG GGA GTC CAG GTG GTG GCC TGT CTG AGT GTT     420
80   Glu Ala Leu Lys Ala Lys Gly Val Gln Val Val Ala Cys Leu Ser Val
81                   65                  70                  75
82
83   AAT GAT GCC TTT GTG ACT GGC GAG TGG GGC CGA GCC CAC AAG GCG GAA     468
84   Asn Asp Ala Phe Val Thr Gly Glu Trp Gly Arg Ala His Lys Ala Glu
85               80                  85                  90
86
87   GGC AAG GTT CGG CTC CTG GCT GAT CCC ACT GGG GCC TTT GGG AAG GAG     516
88   Gly Lys Val Arg Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Glu
89           95                  100                 105
90
91   ACA GAC TTA TTA CTA GAT GAT TCG CTG GTG TCC ATC TTT GGG AAT CGA     564
92   Thr Asp Leu Leu Leu Asp Asp Ser Leu Val Ser Ile Phe Gly Asn Arg
93       110                 115                 120
94
95   CGT CTC AAG AGG TTC TCC ATG GTG GTA CAG GAT GGC ATA GTG AAG GCC     612
96   Arg Leu Lys Arg Phe Ser Met Val Val Gln Asp Gly Ile Val Lys Ala
97   125                 130                 135                 140
98
99   CTG AAT GTG GAA CCA GAT GGC ACA GGC CTC ACC TGC AGC CTG GCA CCC     660
```

RECEIVED

OCT 05 2000

TECH CENTER 1600/2900

INPUT SET: S35961.raw

```
100   Leu Asn Val Glu Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro
101                   145                 150                 155
102
103   AAT ATC ATC TCA CAG CTC TGA GGCCCTGGGC CAGATTACTT CCTCCACCCC          711
104   Asn Ile Ile Ser Gln Leu  *
105               160
106
107   TCCCTATCTC ACCTGCCCAG CCCTGTGCTG GGGCCCTGCA ATTGGAATGT TGGCCAGATT     771
108
109   TCTGCAATAA ACACTTGTGG TTTGCGGAAA AAAA                                  805
110
111
112   (2) INFORMATION FOR SEQ ID NO: 2:
113
114         (i) SEQUENCE CHARACTERISTICS:
115             (A) LENGTH:  162 amino acids
116             (B) TYPE: amino acid
117             (D) TOPOLOGY: linear
118
119       (ii) MOLECULE TYPE: protein
120       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
121
122   Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ala Val Glu Val Phe
123    1               5                 10                  15
124
125   Glu Gly Glu Pro Gly Asn Lys Val Asn Leu Ala Glu Leu Phe Lys Gly
126               20                  25                  30
127
128   Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe Thr Pro Gly Cys
129           35                  40                  45
130
131   Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala Glu Ala Leu Lys
132        50                  55                  60
133
134   Ala Lys Gly Val Gln Val Val Ala Cys Leu Ser Val Asn Asp Ala Phe
135    65                  70                  75                  80
136
137   Val Thr Gly Glu Trp Gly Arg Ala His Lys Ala Glu Gly Lys Val Arg
138               85                  90                  95
139
140   Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Glu Thr Asp Leu Leu
141               100                 105                 110
142
143   Leu Asp Asp Ser Leu Val Ser Ile Phe Gly Asn Arg Arg Leu Lys Arg
144           115                 120                 125
145
146   Phe Ser Met Val Val Gln Asp Gly Ile Val Lys Ala Leu Asn Val Glu
147        130                 135                 140
148
149   Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro Asn Ile Ile Ser
150   145                 150                 155                 160
151
152   Gln Leu
```

INPUT SET: S35961.raw

```
153
154
155     (2) INFORMATION FOR SEQ ID NO: 3:
156
157          (i) SEQUENCE CHARACTERISTICS:
158               (A) LENGTH:  779 amino acids
159               (B) TYPE: nucleic acid
160               (C) STRANDEDNESS: single
161               (D) TOPOLOGY: linear
162
163         (ii) MOLECULE TYPE: DNA (genomic)
164
165        (iii) HYPOTHETICAL: NO
166
167         (iv) ANTI-SENSE: NO
168
169         (vi) ORIGINAL SOURCE:
170               (A) ORGANISM: Rattus Rattus
171
172         (ix) FEATURE:
173               (A) NAME/KEY: CDS
174               (B) LOCATION:136..624
175
176
177         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
178
179     TGCGTCCTAG GCAGCATAGC CGGATCGGTG CTCCGTGCAT CGGCTACTTG GACGTGCGTG     60
180
181     GCAGGCAGAG CAGGCCGGAA AGGAGCAGGT TGGGAGTGTG GTGGGGCCCG CAGCTTCAGC    120
182
183     AGTGCCGCGG TGACTATGGC CCCGATCAAG GTGGGAGACA CCATTCCCTC AGTGGAGGTA    180
184
185     TTTGRAGGGG AACCTGGAAA GAAGGTGAAC TTGGCAGAGC TGTTCAAGGA CAAGAAAGGT    240
186
187     GTTTTGTTTG GAGTCCCTGG GGCATTTACA CCTGGCTGTT CCAAGACCCA TCTGCCTGGG    300
188
189     TTTGTGGAGC AAGCCGGAGC TCYGAAGGCC AAGGGAGCAC AAGTGGTGGC CTGTCTGAGT    360
190
191     GTTAATGATG YCTTCGTGAC TGCAGAGTGG GGTCGAGCCC ACCAGGCAGA AGGCAAGGTT    420
192
193     CAGCTCCTGG CTGACCCCAC TGGAGCTTTT GGAAAGGAGA CAGATTTACT ACTAGATGAT    480
194
195     TCTTTGGTGT CTCTCTTTGG GAATCGTCGG CTAAAAAGGT TCTCCATGGT GATAGACAAG    540
196
197     GGCGTAGTAA AGGCACTGAA CGTGGAGCCG GATGGCACAG GCCTCACCTG CAGCCTGGCC    600
198
199     CCCAACATCC TCTCACAACT CTGAGGCCCT GACCAGAATG TCCTCTGACT CTCCCATCTC    660
200
201     CTCCACCCAG CTCTGGGCCA AAGGCCCAGT ACCTCCTTAC CTGAGGGCCA CTGGAATGGA    720
202
203     ACCTTGACAA TATTTCTGCA ATAAACAGTT TAATTTGTGA AAAAAAAAAA AAAAAAAAAA    780
204
205
```

```
206    (2) INFORMATION FOR SEQ ID NO: 4:
207
208        (i) SEQUENCE CHARACTERISTICS:
209            (A) LENGTH: 162 amino acids
210            (B) TYPE: amino acid
211            (C) STRANDEDNESS: single
212            (D) TOPOLOGY: linear
213
214       (ii) MOLECULE TYPE: peptide
215
216      (iii) HYPOTHETICAL: NO
217
218       (iv) ANTI-SENSE: NO
219
220       (vi) ORIGINAL SOURCE:
221            (A) ORGANISM: Rattus Rattus
222
223       (ix) FEATURE:
224            (A) NAME/KEY: Modified-site
225            (B) LOCATION:17
226            (D) OTHER INFORMATION:/product= "Glu or Gly"
227
228       (ix) FEATURE:
229            (A) NAME/KEY: Modified-site
230            (B) LOCATION:63
231            (D) OTHER INFORMATION:/product= "Leu or Pro"
232
233       (ix) FEATURE:
234            (A) NAME/KEY: Modified-site
235            (B) LOCATION:79
236            (D) OTHER INFORMATION:/product= "Ala or Val"
237
238
239       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
240
241    Met Ala Pro Ile Lys Val Gly Asp Thr Ile Pro Ser Val Glu Val Phe
242    1               5                   10                  15
243
244    Xaa Gly Glu Pro Gly Lys Lys Val Asn Leu Ala Glu Leu Phe Lys Asp
245                20                  25                  30
246
247    Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe Thr Pro Gly Cys
248            35                  40                  45
249
250    Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala Gly Ala Xaa Lys
251        50                  55                  60
252
253    Ala Lys Gly Ala Gln Val Val Ala Cys Leu Ser Val Asn Asp Xaa Phe
254    65                  70                  75                  80
255
256    Val Thr Ala Glu Trp Gly Arg Ala His Gln Ala Glu Gly Lys Val Gln
257                    85                  90                  95
258
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/09/486,167*

INPUT SET: S35961.raw

Line	Error	Original Text
5	Mandatory Value Not Present	(i) APPLICANT:
6	Unknown or Misplaced Identifier	(A) NAME: UNIVERSITE CATHOLIQUE DE LOUVAIN
8	Unknown or Misplaced Identifier	(B) STREET: Place de l' Universite, 1
9	Unknown or Misplaced Identifier	(C) CITY: LOUVAIN-LA-NEUVE
10	Unknown or Misplaced Identifier	(E) COUNTRY: BELGIUM
11	Unknown or Misplaced Identifier	(F) POSTAL CODE (ZIP): B-1348
13	Unknown or Misplaced Identifier	(A) NAME: UNIVERSITE DE MONS-HAINAUT
14	Unknown or Misplaced Identifier	(B) STREET: Place du Parc 20
15	Unknown or Misplaced Identifier	(C) CITY: MONS
16	Unknown or Misplaced Identifier	(E) COUNTRY: BELGIUM
17	Unknown or Misplaced Identifier	(F) POSTAL CODE (ZIP): B-7000
155	Stop Codon at end of sequence removed - no error	(2) INFORMATION FOR SEQ ID NO: 3:

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION   *US/09/486,167*

INPUT SET: S35961.raw

ADDRESSEE
STREET
CITY
STATE
COUNTRY
ZIP
CORRESPONDENCE ADDRESS
APPLICATION NUMBER
FILING DATE
CLASSIFICATION
CURRENT APPLICATION DATA
APPLICATION NUMBER
FILING DATE
PRIOR APPLICATION DATA